

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2002, 14:50:01; Search time 11 Seconds

(without alignments)
867.232 Million cell updates/sec

Title: US-09-856-320A-2_COPY_53_282

Perfect score: 1263
Sequence: 1 KIKGFCECKPHSQPQALF.....GVYTKCKYDVIQETMKNN 230

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	1263	100.0	250	1	KLRB_HUMAN
2	739	58.5	250	1	KLRB_HUMAN
3	689	54.6	260	1	KLRB_HUMAN
4	686	54.3	260	1	KLRB_HUMAN
5	684	54.2	256	1	KLRB_HUMAN
6	683	54.1	260	1	KLRB_HUMAN
7	677.5	53.6	277	1	KLRB_HUMAN
8	649.5	51.4	293	1	KLRB_HUMAN
9	623	49.3	251	1	KLRB_HUMAN
10	610.5	48.3	261	1	KLRB_HUMAN
11	608	48.1	248	1	KLRB_HUMAN
12	605.5	47.9	261	1	KLRB_HUMAN
13	600.5	47.5	261	1	KLRB_HUMAN
14	596.5	47.2	263	1	KLRB_HUMAN
15	596	47.2	248	1	KLRB_HUMAN
16	593	47.0	246	1	KLRB_HUMAN
17	590.5	46.8	261	1	KLRB_HUMAN
18	590	46.7	238	1	KLRB_HUMAN
19	589	46.6	246	1	KLRB_HUMAN
20	586.5	46.4	259	1	KLRB_HUMAN
21	582.5	46.1	261	1	KLRB_HUMAN
22	582.5	46.1	261	1	KLRB_HUMAN
23	579	45.8	231	1	KLRB_HUMAN
24	579	45.8	244	1	KLRB_HUMAN
25	575.5	45.6	246	1	KLRB_HUMAN
26	573	45.4	246	1	KLRB_HUMAN
27	572.5	45.3	259	1	KLRB_HUMAN
28	571.5	45.2	261	1	KLRB_HUMAN
29	570	45.1	247	1	KLRB_HUMAN
30	569.5	45.1	244	1	KLRB_HUMAN
31	569.5	45.1	261	1	KLRB_HUMAN
32	567	44.9	262	1	KLRB_HUMAN
33	566	44.8	243	1	KLRB_HUMAN

34	563	44.6	248	1	TRY1_CHICK
35	561.5	44.5	257	1	KLRB_HUMAN
36	561.5	44.5	261	1	KLRB_HUMAN
37	559	44.3	244	1	TRY2_XENLA
38	559	44.3	248	1	TRY2_CHICK
39	558	44.2	243	1	TRY1_XENLA
40	558	44.2	247	1	TRY2_XENLA
41	556.5	44.1	261	1	TRY2_BOVIN
42	555	43.9	247	1	KLRB_HUMAN
43	554.5	43.9	231	1	TRY2_SALSA
44	554.5	43.9	247	1	TRY2_SALSA
45	554	43.9	258	1	KLRB_HUMAN

ALIGNMENTS

RESULT 1
KLRB_HUMAN STANDARD: PRT: 250 AA.
AC Q9UBX7, 075837; Q9NS65;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Kallikrein II precursor (EC 3.4.21.-) (Hypotensin) (Trypsin-like protease).
GN KLR1 OR PRSS20 OR TLSP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Hippocampus;
RX MEDLINE=98438738; PubMed=9765601;
RA Yoshida S., Taniguchi M., Suemoto T., Oka T., He X.P., Shiosaka S.;
RT "CDNA cloning and expression of a novel serine protease, TLSP.";
RL Biochim. Biophys. Acta 1399:225-228(1998).
[2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Hippocampus, and prostate;
RX MEDLINE=20329229; PubMed=10872828;
RA Mitsui S., Yamada T., Okui A., Komiyama K., Uemura H., Yamaguchi N.;
RT "A novel isoform of a kallikrein-like protease, TLSP/hypotensin, (PRSS20), is expressed in the human brain and prostate.";
RL Biochem. Biophys. Res. Commun. 272:205-211(2000).
[3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20130117; PubMed=10662548;
RA Yousef G.M., Scorrilas A., Diamandis E.P.;
RT "Genomic organization, mapping, tissue expression, and hormonal regulation of trypsin-like serine protease (TLSP PRSS20), a new member of the human kallikrein gene family.";
RL Genomics 63:88-96(2000).
[4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20510030; PubMed=11054574;
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McQuay J., Moss P., Paepker B., Wang K.;
RT "Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";
RL Gene 257:119-130(2000).
[5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Iamerdin J.E., McCready P.M., Skowronski E., Vismathan V., Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S., Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Gaines J., Dargan L., Frankel M., Christensen M., Georgescu A., Avila J., Liu S., Andrease T., Frankel M., Attix C., Amico-Keller G., Coefficient J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B., Arelano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.;
RT "Sequence analysis of chromosome 19q13.4.";

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE-Testis;
 RA Strauberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- PUNCTION: POSSIBLE MULTIFUNCTIONAL PROTEASE. EFFICIENTLY CLEAVES
 CC B2-PHE-ARG-4-METHYLCOMARTL-7-AMIDE, A KALLIKREIN SUBSTRATE, AND
 CC WEAKLY CLEAVES OTHER SUBSTRATES FOR KALLIKREIN AND TRYPSIN.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, SKIN AND PROSTATE. ISOFORM
 CC 1 IS EXPRESSED PREFERENTIALLY IN BRAIN; ISOFORM 2 IN PROSTATE.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB012917; BAA33404.1; ALT_INT.
 DR EMBL: AB013730; BAA86713.1; -.
 DR EMBL: AB041036; BAA96797.1; -.
 DR EMBL: AF164623; AAD47815.1; -.
 DR EMBL: AF243527; AAG33364.1; -.
 DR EMBL: AC011473; AAG3257.1; -.
 DR EMBL: BC022068; AAR22068.1; -.
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.257; -.
 DR Genew: HGNC:6359; KLK11.
 DR MIM: 604434; -.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; TRYP-SPC. 1.
 DR PROSITE: PS050240; TRYPSIN_DOM. 1.
 DR PROSITE: PS00134; TRYPSIN_HIS. 1.
 DR PROSITE: PS00135; TRYPSIN_SER. 1.
 KM Hydrolyase; Serine protease; Glycoprotein; Signal; Zymogen;
 KW Alternative splicing.
 FT SIGNAL 1
 FT PROPEP 19 21 POTENTIAL.
 FT CHAIN 22 250 ACTIVATION PEPTIDE (POTENTIAL).
 FT ACT_SITE 62 62 KALLIKREIN 11.
 FT ACT_SITE 110 110 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 203 203 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULEID 28 163 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULEID 47 63 BY SIMILARITY.
 FT DISULEID 135 237 BY SIMILARITY.
 FT DISULEID 142 209 BY SIMILARITY.
 FT DISULEID 174 188 BY SIMILARITY.
 FT DISULEID 199 224 BY SIMILARITY.
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 1 1 M -> MORLUWLDKMSGGLTAAKPCARRSSPLQAM
 FT (IN ISOFORM 2).
 SQ SEQUENCE 250 AA; 27466 MW; 192D910BBDCD7A56 CRC64;
 Query Match 100.0%; Score 1263; DB 1; Length 250;
 Best Local Similarity 100.0%; Pred. No. 7, 1e-109;
 Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 ECCEQTRATSPHPGPFNNSLPKKDRNDIMLVKMASPVSTMAVRPLTSSRCVACT 120
 DB 81 ECCEQTRATSPHPGPFNNSLPKKDRNDIMLVKMASPVSTMAVRPLTSSRCVACT 140
 QY 121 SCLISGMCSTSPOLRPHTRCANITIEHOKECENYPGNITPTWCASVQEGKSCQ 180
 DB 141 SCLISGMCSTSPOLRPHTRCANITIEHOKECENYPGNITPTWCASVQEGKSCQ 200
 QY 181 GDSGGLPVLCNQSLOGIISWGDPCATRRPGVYTRKYCKYVDWIQETMKN 230
 DB 201 GDSGGLPVLCNQSLOGIISWGDPCATRRPGVYTRKYCKYVDWIQETMKN 250
 RESULT 2
 ID KLK9_HUMAN STANDARD; PRT; 250 AA.
 AC 09UK09;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Kallikrein 9 precursor (EC 3.4.21.-) (Kallikrein-like protein 3) (KLK-
 DE L3).
 GN KLK9.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20118156; PubMed-10652563;
 RA Yousif G.M., Luo L.-Y., Diamandis E.P.;
 RT "Identification of novel human Kallikrein-like genes on chromosome
 RT 19q13.3-q13.4".
 RL Anticancer Res. 19:2843-2852(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20247258; PubMed-10783266;
 RA Yousif G.M., Diamandis E.P.;
 RT "The expanded human Kallikrein gene family: locus characterization and
 RT molecular cloning of a new member, KLK-L3".
 RL Genomics 65:184-194(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20510030; PubMed-11054574;
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
 RA Moss P., Paepel B., Wang K.;
 RT "Sequencing and expression analysis of the serine protease gene
 RT cluster located in chromosome 19q13 region".
 RL Gene 257:119-130(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Lamerdin J.E., McCreedy P.M., Skowronski E., Wiswanathan V.,
 RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Gernes J.,
 RA Dangnanan L., Ertler A., Christensen M., Georgescu A., Avila J., Liu S.,
 RA Andreise T., Frankelm M., Altix C., Amico-Keller G., Coefield J.,
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
 RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
 RA Olsen A.S., Carraro A.V.;
 RT "Sequence analysis of chromosome 19q13.4".
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- TISSUE SPECIFICITY: SKIN, THYMUS, TRACHEA, CEREBELLUM AND SPINAL
 CC CORD.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```
CC -----
DR EMBL; AF135026; AAD26427.2; -
DR EMBL; AF243527; AAC33362.1; -
DR EMBL; AC011473; AAG23255.1; -
DR HSSP; P00763; IDPO.
DR MEROPS; S01.307; -.
DR Geneweb; HGNC:6370; KIK9.
DR MIM; 605504; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; Trysin; 1.
DR PRINTS; PR00722; CHYMOTRYPIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KM Hydroxylase; Serine protease; Glycoprotein; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 250 KALLIKREIN 9.
FT ACT_SITE 63 63 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 111 111 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 204 204 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 29 164 BY SIMILARITY.
FT DISULFID 48 64 BY SIMILARITY.
FT DISULFID 136 238 BY SIMILARITY.
FT DISULFID 143 210 BY SIMILARITY.
FT DISULFID 175 189 BY SIMILARITY.
FT DISULFID 200 225 BY SIMILARITY.
FT CARBOHYD 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 211 211 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 250 AA; 27512 MW; F2785245B03E98B CRC64;

Query Match      58.5%; Score 739; DB 1; Length 250;
Best Local Similarity 58.1%; Pred. No. 8,8e-61;
Matches 133; Conservative 35; Mismatches 61; Indels 0; Gaps 0;

QY 1 RINGECKPPHSQPOMALFEKTRLLGATLIAPRWLLTAHCLKPRIYVHLGHNLKE 60
   +-+ +++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++
Db 22 RAIGAECEBPSPSOQAOLFLHTRLFCAGATLISDRWLLTAAICRKPYLMVRIGEHLWM 81

QY 61 EGCEBTATRTATSEFPHPGNNSLPNDHRNDDIMLYMAASPVSTMAVPRULTSSRCVTAGT 120
   +++ ++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++
Db 82 EGPEQLFVYTDFPFPHGKNKDLSANDHDDMLIRLPQARISPAVQPLNTSQTCTVSPEM 141

QY 121 SCLISGMGSTSPQRLRHPTLRCAANTITIEHKOCENAPNTIDPMVCAVSGEGKDSQ 180
   ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++
Db 142 QCLISGMCAVSPKALFVTLQCANTSLLENKLCHMAIAPHGISDMLCAGLEGGRGSQ 201

QY 181 GDSGGPLVCNOSLOGIISMGDPCAITKKPGVYTYVCKIVDMIQETMKN 229
   ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++
Db 202 GDSGGPLVCNGTLGVSGAEPCSRPRRPAYTVSVCHLDWIQEIEMEN 250

RESULT 3
NRPN_RAT
ID NRPN_RAT STANDARD; PRT; 260 AA.
AC 086780;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUN-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neutropin precursor (EC 3.4.21.-) (NP) (Kallikrein 8) (Brain serine
  De protease 1)
OS KIK8 OR PRSS19 OR NRPN OR BSP1.
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer; TISSUE=Brain;
MD MEDLINE=98389725; PubMed=9722524;
NA Davies B.J., Pickard B.S., Steel M., Morris R.G.M., Lathé R.;
```

[illegible]

DE Neuropilin precursor (EC 3.4.21.-) (NP) (Kallikrein 8).
 GN KLR8 OR PRSS19 OR NRPN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathia; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-Hippocampus;
 RX MEDLINE=95348817; PubMed=7621317;
 RA Chen Z.-L., Yoshida S., Kato K., Momota Y., Suzuki J., Tanaka T.,
 RA Ito J., Nishino H., Aimoto S., Kiyama H., Shiosaka S.;
 RT "Expression and activity-dependent changes of a novel limbic-serine
 RT protease gene in the hippocampus.";
 RL J. Neurosci. 15:5088-5097(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yoshida S., Hirata A., Inoue N., Shiosaka S.;
 RT "Cloning and assignment of mouse neuropilin gene, Prss19 to chromosome
 RT 7B4.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF N-TERMINUS AND CHARACTERIZATION.
 RC STRAIN-BALB/C; TISSUE-Brain;
 RX MEDLINE=98225202; PubMed=9536608;
 RA Shimizu C., Yoshida S., Shibata M., Kato K., Momota Y., Matsumoto K.,
 RA Shiosaka T., Madorikawa R., Kamachi T., Kawabe A., Shiosaka S.;
 RT "Characterization of recombinant and brain neuropilin, a
 RT plasticity-related serine protease.";
 RL J. Biol. Chem. 273:11189-11196(1998).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 33-257.
 RC TISSUE-Hippocampus;
 RX MEDLINE=99134351; PubMed=9933620;
 RA Kishi T., Kato M., Shimizu T., Kato K., Matsumoto K., Yoshida S.,
 RA Shiosaka S., Hakoshima T.;
 RT "Crystal structure of neuropilin, a hippocampal protease involved in
 RT kindling epileptogenesis.";
 RL J. Biol. Chem. 274:4220-4224(1999).
 CC -1- FUNCTION: SUGGESTED TO BE INVOLVED IN KINDLING EPILEPTOGENESIS AND
 CC HIPPOCAMPAL PLASTICITY. HAS A STRONG PROTEOLYTIC ACTIVITY AGAINST
 CC FIBRONECTIN.
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
 CC -1- ENZYME REGULATION: STRONGLY INHIBITED BY DIISOPROPYL
 CC FLUOROPHOSPHATE, LEUPEPTIN AND (4-AMIDINOPHENYL)METHANESULFONYL 1-
 CC FLUORIDE.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE LIMBIC SYSTEM OF
 CC MOUSE BRAIN AND IS LOCALIZED AT HIGHEST CONCENTRATION IN PYRAMIDAL
 CC NEURONS OF THE HIPPOCAMPAL CA1-3 SUBFIELDS.
 CC -1- MASS SPECTROMETRY: MW=26613; METHOD-MALDI; RANGE=29-260.
 CC -1- MASS SPECTROMETRY: MW=26229; METHOD-MALDI; RANGE=33-260.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUPERFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D30785; BAA06451.1; -
 DR EMBL: AB032302; BAA92435.1; -
 DR PDB: 1NPM; 23-MAR-99.
 DR MEROPS: S01.244; -
 DR MGD: MGI:892018; KIX8.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; Tryp_SPC; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.

DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KM Hydrolyse: Serine protease; Glycoprotein; zymogen; Signal;
 KW 3D-structure.
 FT SIGNAL 1 28 POTENTIAL.
 FT PROPER 29 32
 FT CHAIN 33 260
 FT ACT_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 39 173
 FT DISULFID 58 74
 FT DISULFID 145 246
 FT DISULFID 152 218
 FT DISULFID 184 198
 FT DISULFID 208 233
 FT CARBOHYD 110 110
 SQ SEQUENCE 260 AA; 28523 MW; B5F6F6BE37CDD0E CRC64; N-LINKED (GLCNAC...); (POTENTIAL).
 Query Match 54.3%; Score 686; DB 1; Length 260;
 Best Local Similarity 52.8%; Pred. No. 6,7e-56;
 Matches 121; Conservative 36; Mismatches 70; Indels 2; Gaps 2;
 QY 1 RIINGECKRHSPWQAALPEKTLICGATLILAPRLTLTAHCLKPRYIYHLOHNLQKE 60
 DB 32 KILGRCRIPHSOPWQALFGEGLIGVLYGDRWVLTAAHCKKORYSVRLGDSLSQR 91
 QY 61 EGCROTATATSPFPHPFNNSLPNKHDRNDIMLVKMASPSITWAVPLTSSRCVAGT 120
 DB 92 DQPEQELQVQAOISHPICYNNSNP-EDISHDMLIRLNSANLGDKVAPVOLANICPRVQ 150
 QY 121 SCLISGWSSTSPQLRPLHRLCANITIIIEHOKENAVPQNTDTWVCASVOEGKDSQ 180
 DB 151 KCLISGWTYTPSPENPNTLNCALVEKIVSONKCEKRAYPKITGEMVACG-SSNGADTQ 209
 QY 181 GDSGRLVCNOSLOGIISWGDPCATIRKPGYTKVKYKIDWIOETKKN 229
 DB 210 GDSGRLVCNOMLOGITISWSDPCGKPEKPGYTKYKIRYTTWIKTKMDN 258
 RESULT 5
 KLR8_HUMAN STANDARD; PRT; 256 AA.
 ID KLR8_HUMAN
 AC Q9H2R5; Q9H2R6; Q9H2R4; Q9H2R3; Q9HBG9; Q15358;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Kallikrein 15 precursor (EC 3.4.21.-) (ACO protease).
 GN KLR15.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RX PubMed=11010966;
 RA Yousef G.M., Scoriilas A., Jung K., Ashworth L.K., Diamandis E.P.;
 RT "Molecular cloning of the human kallikrein 15 gene (KLR15). Up-
 RT regulation in prostate cancer.";
 RL J. Biol. Chem. 276:53-61(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20510030; PubMed=11054574;
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
 RA Moss P., Paepker B., Wang K.;
 RT "Sequencing and expression analysis of the serine protease gene
 RT cluster located in chromosome 19q13 region.";
 RL Gene 257:119-130(2000).
 RN [3]
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE=94289486; PubMed=80187726;
 RA Dhanich M.E., Spleess M.;

RT "A novel serine proteinase-like sequence from human brain.";
 RL Biochim. Biophys. Acta 1218:225-228 (1994).
 CC -1- FUNCTION: Protease whose physiological substrate is not yet known.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; 1 (shown here), 2, 3 and 4; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Highest expression in the thyroid gland. Also
 CC expressed in the prostate, salivary, and adrenal glands and in the
 CC colon testis and kidney.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AF242195; AAG09469.1; -;
 DR EMBL: AF242195; AAG09470.1; -;
 DR EMBL: AF242195; AAG09471.1; -;
 DR EMBL: AF242195; AAG09472.1; -;
 DR EMBL: AF243527; AAG33354.1; -;
 DR EMBL: X75363; CAAS3145.1; ALT_SEQ.
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.081; -;
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser.protease_Try.
 DR Pfam: PF00089; trypsin.1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; TRY_PSPC.1.
 DR PROSITE: PS02040; TRYPSIN_DOM.1.
 DR PROSITE: PS00134; TRYPSIN_HIS.1.
 DR PROSITE: PS00135; TRYPSIN_SER; FALSE_NEG.
 KM Hydroxylase, Serine protease; Glycoprotein; Signal; Zymogen;
 KW Alternative splicing.
 FT SIGNAL 1 16
 FT PROPEP 17 21
 FT CHAIN 22 256
 FT ACT_SITE 62 62
 FT ACT_SITE 106 106
 FT ACT_SITE 209 209
 FT CARBOHYD 171 171
 FT CARBOHYD 232 232
 FT CARBOHYD 232 232
 FT CARBOHYD 206 206
 FT VARSPPLIC 122 256
 FT VARSPPLIC 161 161
 FT VARSPPLIC 162 256
 FT VARSPPLIC 147 160
 FT CONFLICT 147 160
 SQ SEQUENCE 256 AA; 28087 MW; B5EBF8D6022786B5 CRC64;
 Query Match 54.2%; Score 684; DB 1; Length 256;
 Best Local Similarity 51.7%; Pred. No. 1e-55;
 Matches 124; Conservative 37; Mismatches 65; Indels 14; Gaps 3;

RESULT 6
 KLR8_HUMAN STANDARD; PRT; 260 AA.
 AC 060259; 09U047; 09HC83; 09U119;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neuropilin precursor (EC 3.4.21.-) (NP) (Kallikrein 8) (Ovasin) (Serine
 DE protease TADG-14) (Tumor-associated differentially expressed gene-14
 DE protein).
 DE
 GN KLR8 OR PRSS19 OR TADG14 OR NRP.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE-Hippocampus.
 RA MEDLINE=98372070; PubMed=9714609;
 RT Yoshida S., Taniguchi M., Hirata A., Shiosaka S.;
 RL "Sequence analysis and expression of human neuropilin cDNA and gene.";
 RN gene 213:9-16(1998).
 [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE-Brain;
 RA MEDLINE=99203457; PubMed=10102990;
 RT Mitsui S., Tsuruoka N., Yamashiro K., Nakazato H., Yamaguchi N.;
 RL "A novel form of human neuropilin, a brain-related serine protease, is
 RT generated by alternative splicing and is expressed preferentially in
 RL human adult brain.";
 RL Eur. J. Biochem. 260:627-634(1999).
 [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE-Ovary;
 RA MEDLINE=99413504; PubMed=10485494;
 RT Underwood L.J., Tanihito H., Wang Y., Shigemasa K., Parmley T.H.,
 RA O'Brien T.J.;
 RL "Cloning of tumor-associated differentially expressed gene-14, a novel
 RT serine protease overexpressed by ovarian carcinoma.";
 RL Cancer Res. 59:4435-4439(1999).
 [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Gan L., Gellinas R., Gown A.M., Moss P., Smith R., Wang K.;
 RT "Molecular cloning and characterization of a novel serine protease,
 RT ovasin, a potential molecular marker for ovarian carcinomas.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=20510030; PubMed=11054574;
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
 RA Moss P., Paepert B., Wang K.;
 RT "Sequencing and expression analysis of the serine protease gene
 RT cluster located in chromosome 19q13 region.";
 RL gene 257:119-130(2000).
 [6]
 RP SEQUENCE OF 1-164 FROM N.A. (ISOFORM 1).
 RA Lamerding J.E., McCready P.M., Skowronski E., Viswanathan V.,
 RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilaagen S.,
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Gaines J.,
 RA Danganan L., Trankheim M., Attix C., Amico-Keller G., Coefield J.,
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
 RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
 RA Olsen A.S., Carraro A.V.;
 RT "Sequence analysis of chromosome 19q13.4.";
 RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RL
 CC -1- FUNCTION: SUGGESTED TO BE INVOLVED IN KINDLING EPILEPTOGENESIS AND
 CC HIPPOCAMPAL PLASTICITY.
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: ISOFORM 1 IS PREDOMINANTLY EXPRESSED IN THE

```

CC PANCREAS WHITE ISOFORM 2 IS EXPRESSED IN ADULT BRAIN AND
CC HIPPOCAMPUS. BOTH FORMS ARE ALSO FOUND IN FETAL BRAIN AND
CC PLACENTA. NOT DETECTED IN KIDNEY, SPLEEN, LIVER AND LUNG.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB009849; BAA28673.1; -
DR EMBL; AB012761; BAA28676.1; -
DR EMBL; AB010780; BAA8684.1; -
DR EMBL; AB008390; BAA82665.1; -
DR EMBL; AB008927; BAA82666.1; -
DR EMBL; AF055982; AAD56050.1; -
DR EMBL; AF095742; AAD5979.1; -
DR EMBL; AF095743; AAD29574.1; -
DR EMBL; AF243527; AAG33361.1; -
DR EMBL; AC011473; AAG23254.1; -
DR HSSP; 061955; INPM.
DR MEROPS; S01.244; -
DR Genew; HGNC:6369; KLK8.
DR MIM; 605644; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin.1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease; Glycoprotein; zymogen; signal;
KW Alternative splicing.
KM SIGNAL 1 28
FT 1 28
FT PROPE 29 32
FT CHAIN 33 260
FT ACT_SITE 73 73
FT ACT_SITE 120 120
FT ACT_SITE 212 212
FT DISULFID 39 173
FT DISULFID 58 74
FT DISULFID 145 246
FT DISULFID 152 218
FT DISULFID 184 198
FT DISULFID 208 233
FT CARBOHYD 110 110
FT VARSPLIC 23 23
FT SEQUENCE 260 AA; 28048 MW; EF4395B8C3B660 CRC64;
SQ
Query Match 54.1%; Score 683; DB 1; Length 260;
Best Local Similarity 51.6%; Pred. NO. 1.3e-55;
Matches 116; Conservative 43; Mismatches 64; Indels 2; Gaps 2;

```

```

RESULT 7
ID KIKD_HUMAN STANDARD; PRT; 277 AA.
AC Q9UKR3; Q9Y433;
DT 16-OCT-2001 (rel. 40, Created)
DT 16-OCT-2001 (rel. 40, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Kallikrein 13 precursor (EC 3.4.21.-) (Kallikrein-like protein 4)
DE (KIK-14).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20229789; PubMed-10766816;
RA Yousef G.M., Chang A., Diamandis E.P.;
RT "Identification and characterization of KLK-14, a new kallikrein-like
RT gene that appears to be down-regulated in breast cancer tissues.";
RL N. Engl. J. Biol. Chem. 275:11891-11898(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Gaines J.,
RA Dargatzis L., Erlar A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Andrade T., Frankel M., Altix C., Amico-Keller G., Coiffield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrano A.V.;
RT "Sequence analysis of chromosome 19q13.4.";
RL Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-180 FROM N.A.
RC TISSUE=uterus;
RA Ansoorge W., Wilkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN PROSTATE, BREAST, TESTIS AND
CC SALIVARY GLAND.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF135024; AAD26425.2; -
DR EMBL; AC011473; AAG23259.1; -
DR EMBL; AL050220; CAB3320.1; ALT_INIT.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.306; -
DR Genew; HGNC:6361; KLK13.
DR MIM; 605505; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin.1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Glycoprotein; signal.
FT SIGNAL 1 16
FT CHAIN 17 277
FT ACT_SITE 76 76
FT ACT_SITE 124 124
FT ACT_SITE 124 124
FT CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CHARGE RELAY SYSTEM (BY SIMILARITY).

```

```

FT ACT_SITE 218 218 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 42 178 BY SIMILARITY.
FT DISULFID 61 77 BY SIMILARITY.
FT DISULFID 157 224 BY SIMILARITY.
FT DISULFID 189 203 BY SIMILARITY.
FT DISULFID 214 239 BY SIMILARITY.
FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 170 180 UNYPKLOCAN -> GMPHPRMEAP (IN REF. 3).
SQ SEQUENCE 277 AA: 30570 MW: BA8A9EBDCFB5D542 CRC64;

Query Match 53.6%; Score 677.5; DB 1; Length 277;
Best Local Similarity 53.3%; Pred. No. 4,3e-55;
Matches 120; Conservative 41; Mismatches 63; Indels 1; Gaps 1;

OY 5 GFCEKPHSQWQALFEKTRLLCGATLLAPRWLTAAHCKLPRIYVHLGOHNLKEGCE 64
DB 39 GYCFPHSQWQALVQGLLGGVLYHNRKWTAAHCKLEGKLYLGNHALGRVAGE 98
OY 65 QTRATESFPHGPNNSLPKNDHNDIMLVKMASPVSTMAVRPLTSL-SRCYTAGTSC 123
DB 99 QREVVHSHIPHEXRRSPTHLNDHDIHLLELQSPVQLTGVIQTLPLSHNNRLTPGTCR 158
OY 124 ISGMSYSSQQLRPHTRLCANTITIEHOKCENAPGNTIDTVKASVQEGKDSQCGDS 183
DB 159 VSGMGTTSPOVNYPKTLCQANIQLRSEDECRQYPERKIDNMICAGTEKSGDCSGDS 218
OY 184 GGPVACNOSLOGIISWGDPCATRRKPGVYTKYKVDWIOETMK 228
DB 219 GGPVACNKTIGYISWGDPCGPDREGVYTRYSRYLWIRETR 263

RESULT 8
KLKE_HUMAN
ID KLKE_HUMAN STANDARD; PRT; 293 AA.
AC 09Y337: 09HGB8.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Kallikrein 5 precursor (BC 3.4.21.-) (Stratum corneum tryptic enzyme)
DE (Kallikrein-like protein 2) (KLK-L2).
GN KLK5 OR SCTE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stratum corneum;
RA MEDLINE=9445563; PubMed=10514489;
RA Bratsand M., Egelrud T.;
RT "Purification, molecular cloning, and expression of a human stratum
RT corneum trypsin-like serine protease with possible function in
RT desquamation."
RL J. Biol. Chem. 274:30033-30040(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=20118156; PubMed=10652563;
RA Yousef G.M., Luo L.-Y., Diamandis E.P.;
RT "Identification of novel human kallikrein-like genes on chromosome
RT 19q13.3-q13.4."
RL Anticancer Res. 19:2843-2852(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE=20510030; PubMed=11054574;
RA Gao L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCaig J.,
RA Moss P., Paepel B., Wang K.;
RT "Sequencing and expression analysis of the serine protease gene
RT cluster located in chromosome 19q13 region."
RL Gene 257:119-130(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=ovary;

```

```

RA Strausberg R.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY BE INVOLVED IN DESQUAMATION.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SKIN, BREAST, BRAIN AND TESTIS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF168768; AAF03101.1; -
DR EMBL: AF135028; AAD26429.1; -
DR EMBL: AF243527; AAG33358.1; -
DR EMBL: BC008036; AAO08036.1; -
DR HSSP: P00763; IDPO.
DR MEROPS: S01.017; -.
DR GeneW: HGNC:6366; KLK5.
DR MIM: 605643;
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Tryp.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Glycoprotein; signal.
FT SIGNAL 1 22
FT CHAIN 23 293 KALLIKREIN 5.
FT ACT_SITE 108 108 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 153 153 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 245 245 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 73 206 BY SIMILARITY.
FT DISULFID 93 109 BY SIMILARITY.
FT DISULFID 178 279 BY SIMILARITY.
FT DISULFID 185 251 BY SIMILARITY.
FT DISULFID 217 231 BY SIMILARITY.
FT DISULFID 241 266 BY SIMILARITY.
FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 25 36 MISSING (IN REF. 3).
SQ SEQUENCE 293 AA: 32020 MW: D92C92F5609E5946 CRC64;

Query Match 51.4%; Score 649.5; DB 1; Length 293;
Best Local Similarity 50.9%; Pred. No. 1.7e-52;
Matches 118; Conservative 40; Mismatches 67; Indels 7; Gaps 4;

OY 1 RIIGFCEKPHSQWQALFEKTRLLCGATLLAPRWLTAAHCKLPRIYVHLGOHNLK 59
DB 66 RIINGSDCMHQRPWQALLLRPNQLYCGAVLVHPQLTLAAHCKKVFVRIGHSLSP 125
OY 60 -EEGCEQTRATSFPHGPNNSLPKNDHNDIMLVKMASPVSTMAVRPLTSLSSRCY 118
DB 126 VYESGQMGVGSIRPGYS---HPGSHNDMLKLRIRIRPTKDVRINVSHPSPA 181
OY 119 GFSCLISGMSSTSPQRLPHTRLCANTITIEHOKCENAPGNTIDTVKASVQEGKDS 178
DB 182 GTGCLVSGMGTTSPOVNYPKTLCQANISVLSQRCRDAPRQIDDTMFCAG-DKAGRDS 240
OY 179 CGGDSGGPVLVACNOSLOGIISWGDPCATRRKPGVYTKYKVDWIOETMKNN 230
DB 241 CGGDSGGPVLVACNOSLOGIISWGDPCARPRRPGVYTMCKFTKWIQETIQAN 292

RESULT 9
KLKE_HUMAN

```

ID	KLK6_HUMAN	STANDARD:	PRT:	251 AA.
AC	Q9P0G3:			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Kallikrein 14 precursor (EC 3.4.21.-) (Kallikrein-like protein 6) (KLK-L6).			
DE	KLK14 OR KLK16.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Yousef G.M., Diamandis E.P.;			
RT	"Molecular characterization, mapping, and tissue expression of KLK16,			
RT	a hormonally regulated kallikrein-like gene.";			
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.			
RP	[2]			
RA	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.			
RA	Hooper J.D., Bul L.T., Rae F.K., Harvey T.J., Myers S.A.,			
RA	Ashworth L.K., Clements J.A.;			
RT	"Identification and characterization of KLK14, a novel kallikrein			
RT	serine protease gene located on human chromosome 19q13.4 and expressed			
RT	in prostate and skeletal muscle.";			
RL	Genomics 73:117-122(2001).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,			
RA	Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,			
RA	Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Ganes J.,			
RA	Danganan L., Erlar A., Christensen M., Georgescu A., Avila J., Liu S.,			
RA	Andreise S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,			
RA	Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,			
RA	Olsen A.S., Carraro A.V.;			
RT	"Sequence analysis of chromosome 19q13.4.";			
RL	Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	TISSUE SPECIFICITY.			
RX	MEDLINE=20545474; PubMed=10963073;			
RA	Harvey T.J., Hooper J.D., Myers S.A., Stephenson S.A., Ashworth L.K.,			
RA	Clements J.A.;			
RT	"Tissue-specific expression patterns and fine mapping of the human			
RT	kallikrein (KLK) locus on proximal 19q13.4.";			
RL	J. Biol. Chem. 275:37397-37406(2000).			
CC	-1- SUBCELLULAR LOCATION: Secreted (Probable).			
CC	-1- TISSUE SPECIFICITY: HIGH EXPRESSION IN BRAIN, BONE MARROW AND			
CC	FETAL LIVER. ALSO EXPRESSED IN LIVER, PANCREAS, FETAL SPLEEN,			
CC	PROSTATE AND SKELETAL MUSCLE.			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, KALLIKREIN SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/ ,			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL, AF161221; AAD50773.2; -			
DR	EMBL, AF283669; AAK48523.1; -			
DR	EMBL, AF283670; AAK48524.1; -			
DR	EMBL, AC011473; AAG3260.1; -			
DR	HSSP; P00763; IDPO.			
DR	MEROPS; S01.029; -			
DR	Genew; HGNC:6362; KLK14.			
DR	MIM; 606135; -			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR001254; Ser_protease_Try.			
DR	Pfam; PF00089; trypsin; 1			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM00020; Tryp_Spc; 1.			

DR	PROSITE	PS00240:	TRYPSIN_DOM.1.
DR	PROSITE	PS00134:	TRYPSIN_HIS.1.
DR	PROSITE	PS00135:	TRYPSIN_SER.1.
KW	Hydrolase;	Serine protease;	Signal; Zymogen.
FT	SIGNAL	1	POTENTIAL.
FT	PROPEP	19	ACTIVATION PEPTIDE (POTENTIAL).
FT	CHAIN	25	KALLIKREIN 14.
FT	ACI_SITE	67	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACI_SITE	111	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACI_SITE	204	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID	31	BY SIMILARITY.
FT	DISULFID	52	BY SIMILARITY.
FT	DISULFID	143	BY SIMILARITY.
FT	DISULFID	175	BY SIMILARITY.
FT	DISULFID	200	BY SIMILARITY.
SO	SEQUENCE	251 AA;	27452 MW; 9087953BAFAED25 CRC64;
<hr/>			
Query Match 49.3%; Score 623; DB 1; Length 251;			
Best Local Similarity 48.9%; Pred. No. 3.9e-50;			
Matches 113; Conservative 38; Mismatches 74; Indels 6; Gaps			
<hr/>			
QY	1	RIRKGFCKPHSPQQAALF--EKTRLLCGATLAPRLMLTAHQAKLPRIYLHGLGHNLIQ	58
Db	24	KITGGHTCTRSSQPQAALAGPRRRRLTCGALLSGGVITAHCGRILLVQALGKNHLR	83
QY	59	KEEGCEQTRTATESEFPHPGEFNNSLPNKDHRNDIMLVKMASFVSITMAVRPLTLSSRCVTA	118
Db	84	RWEATOQLRVAROVTHPENYN---SRFDNDLMILLOQPARIGRAVRPIEVTQAQASP	138
QY	119	GTSCLSIWGSSSTSPQLAPTHLCANTITEEHOKCEANVAGNTIDPMWCASVOEGKDS	178
Db	140	GTSICVSWGTTISSIARYPASLDCCVNINISPDCECAAPRTITPGVCAGVPQGKDS	199
QY	179	CQGSQGGVLVNCOSLOGIISMGDCPCATTARKPGVYTKCKRYVDIIQETMKN	229
Db	200	CQGSQGGVLVCRGDLQGLVGSMERCALPGVPGYYTNCKRSMIEETMRD	250
<hr/>			
RESULT 10	ID	KLK1_RAT	STANDARD; PRT; 261 AA.
AC	21-JUL-1986	(Rel. 01, Created)	
DT	01-JUN-1994	(Rel. 29, Last sequence update)	
DT	15-DEC-1998	(Rel. 37, Last annotation update)	
DE	Glandular kallikrein, pancreatic-1 precursor (EC 3.4.21.35) (tissue		
DE	kallikrein) (PS kallikrein) (RGR-1).		
GN	KLK1 OR KLK-1.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus		
OX	NCBI_Taxid=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=83117659; PubMed=6961406;		
RA	Swift G.H., Dagorn J.-C., Ashley P.L., Cummings S.W., McDonald R.J.;		
RT	"rat pancreatic kallikrein mRNA: nucleotide sequence and amino acid		
RT	sequence of the encoded preproenzyme.";		
RT	Proc. Natl. Acad. Sci. U.S.A. 79:7263-7267(1982).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Submaxillary gland;		
RX	MEDLINE=86051477; PubMed=29964455;		
RA	Ashley P.L., MacDonald R.J.;		
RT	"kallikrein-related mRNAs of the rat submaxillary gland: nucleotide		
RT	sequences of four distinct types including tonin.";		
RL	Biochemistry 24:4512-4520(1985).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Kidney;		
RX	MEDLINE=89327211; PubMed=2755879;		
RA	Inoue H., Fukui K., Miyake Y.;		
RT	"identification and structure of the rat true tissue kallikrein gene		

RT expressed in the kidney.";
 RN J. Biochem. 105:834-840(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89214217; PubMed=2708383;
 RA Wines D.R., Brady J.M., Pritchett D.B., Roberts J.L., Macdonald R.J.:
 RT "Organization and expression of the rat kallikrein gene family.";
 RL J. Biol. Chem. 264:7653-7662(1989).
 RN [5]
 RN SEQUENCE OF 48-261 FROM N.A.
 RX MEDLINE=86131678; PubMed=3004582;
 RA Gerald W.L., Chao J., Chao L.:
 RT "Immunological identification of rat tissue kallikrein cDNA and
 RT characterization of the kallikrein gene family.";
 RL Biochem. Biophys. Acta 866:1-14(1986).
 CC -1- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS
 CC IN KININOGEN TO RELEASE LYS-BRADYKININ.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage of Arg-Lys bonds in
 CC small molecule substrates. Highly selective action to release
 CC kallidin (Lysyl-bradykinin) from kininogen involves hydrolysis of
 CC Met-Lys or Leu-Lys.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL J00758; NOT_ANNOTATED_CDS.
 DR EMBL M11563; AAA41464.1; ALT_INIT.
 DR EMBL M23876; AAA41462.1; -.
 DR EMBL M23874; AAA41462.1; JOINED.
 DR EMBL M23875; AAA41462.1; JOINED.
 DR EMBL D00448; BAA00346.1; ALT_INIT.
 DR EMBL D00446; BAA00346.1; JOINED.
 DR EMBL D00447; BAA00346.1; JOINED.
 DR EMBL X03560; CAA27247.1; -.
 DR PIR A00944; KORTP.
 DR PIR A23863; A23863.
 DR PIR JX0073; JX0073.
 DR PIR A33359; A33359.
 DR HSSP; P00757; ISGF.
 DR MEROPS; S01.405; -.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser.protease_Try.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRY_PSPC; 1.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KM Hydrolyase; Serine protease; Multigene family; Zymogen;
 KM Signal.
 FT SIGNAL 1 18
 FT PROBAB.
 FT PROPEP 19 24
 FT CHAIN 25 261
 FT CHAIN 111
 FT CHAIN 112
 FT ACT_SITE 65 66
 FT ACT_SITE 120 120
 FT ACT_SITE 213 213
 FT DISULFID 31 173
 FT DISULFID 50 66
 FT DISULFID 152 219
 FT DISULFID 184 198
 FT DISULFID 209 234
 FT CARBOHYD 108 108
 FT SEQUENCE 261 AA; 28852 MW; F2F99C0227A7882B CRC64;
 Query Match 48.3%; Score 610.5; DB 1; Length 261;

Best Local Similarity 45.6%; Pred. No. 5.7e-49;
 Matches 108; Conservative 43; Mismatches 79; Indels 7; Gaps 1;
 OY 1 RIHKFECKPHSOPMOALPEKTRFLCGATLLIAPRWLLFAHCKAPRYIYLGHONLQKE 60
 DB 24 RVGVCYCEMNSOPMOVAAYFYFGEYLCGVILIDPSWITPAHCACTDNYQVWLGNNLYED 83
 OY 61 EGGEQRTATESPPHPGFNSNL-----PNKDRNDIMLVKMASPVSIWAVRPLTSS 113
 DB 84 EPPAQHRLVQSPPHPGFNDLIMNHTROPQDDYSNDMLHLISQPADITDGVAVIDLPI 143
 OY 114 RCTYAGTSCILSGWSTSSPOLRPHTLRCANITIIHQKCNENYRGENTIDTWACASVQE 173
 DB 144 EEPVAGSTCLASGWSITPPGLELSDLOCVNIDILSNKCEVNAKKEVTDMLCAGEMD 203
 OY 174 GGDSCGDSGSGPLVYCNQSLGIIISMGDDCAITRKGVYTKCKYVDWIOETPKNN 230
 DB 204 GGDYCKGDSGGLICNGVLQGITSWGPNPCGEKRGITKIKTPWIKREYKEN 260
 RESULT 11
 ID KIKC_HUMAN STANDARD; PRT; 248 AA.
 AC Q9UKR0; Q9UKR1.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Kallikrein 12 precursor (EC 3.4.21.-) (Kallikrein-like protein 5)
 DE (KIK-15).
 GN KIK12 OR KIKL5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=20118156; PubMed=10652563;
 RA Youssef G.M., Luo L.-Y., Diamandis E.P.:
 RT "Identification of novel human kallikrein-like genes on chromosome
 RT 19q13.3-q13.4.";
 RL 19q13.3-q13.4.";
 RL Anticancer Res. 19:2843-2852(1999).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RA Youssef G.M., Magklara A., Scorilas A., Diamandis E.P.:
 RT "Cloning of new alternatively spliced forms of the kallikrein-like
 RT gene 5 (KIK-15)."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=20510030; PubMed=11054574;
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
 RA Moss P., Pieper B., Wang K.:
 RT "Sequencing and expression analysis of the serine protease gene
 RT cluster located in chromosome 19q13 region.";
 RL gene 257:119-130(2000).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
 RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillagen S.,
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Gaines J.,
 RA Dangnan L., Erlar A., Christensen M., Georgescu A., Avila J., Liu S.,
 RA Andreise T., Frankheim M., Atlix C., Amico-Keller G., Coetfield B.,
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
 RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
 RA Olsen A.S., Carraro A.V.:
 RT "Sequence analysis of chromosome 19q13.4.";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

QY 1 RIIFGECKRHSQPWQALFEKTRLLCGATLIPRLMLTAHCLKPRYIVHGOHNLQKE 60
 DB 24 RVIIGGCKEKNSSQPMWVAAYRYQYLOGVLLPNNVLAFAHCYDDMYKWLCKNNLFXD 83
 QY 61 ECGEOTRTATESRPHFGFNNSL-----DNKRDIMLVKASPVSTMAVRPLTLSS 113
 DB 84 EPPAQRHVLVSQSPHPYKPFELMKNRKRGDDSHNMLHLHSQPDITDGVKVDLPT 143
 QY 114 RCVYAGTSCILSGWGTSSPOLRLPRTLRCANTTIEHOKCEAAYPGNTDPMVCAVQOE 173
 DB 144 EEPKVGSTCLASGWSGSKPLPMEFPDLDQCVNHLHLSNECIRAYKREKVTIDMILCAGELE 203
 QY 174 GGDSCQSGSGGGLPVNCSQIGIISWGDPCAITRRPGVTVKCYVYDIQETMKN 230
 DB 204 GGDTCYGDGSGGLDGVLOGITSWGHPGCAKTMNPAITKLTFTSWIKWEYKEN 260

RESULT 13
 KLR3_MOUSE STANDARD: PRT: 261 AA.
 AC P00756;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glandular kallikrein K3 precursor (EC 3.4.21.35) (tissue kallikrein)
 DE (mGK-3) (7S nerve growth factor gamma chain) (gamma-NGF).
 GN KLR3 OR KLR-3 OR NGFG.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RN SEQUENCE FROM N.A.
 RX MEDLINE=85076169; PubMed=6548955;
 RA Ullrich A., Gray A., Wood W.I., Hayflick J., Seeburg P.H.;
 RT "Isolation of a cDNA clone coding for the gamma-subunit of mouse
 RT nerve growth factor using a high-stringency selection procedure";
 RL DNA 3:387-392(1984).
 RN (2)
 RN SEQUENCE FROM N.A.
 RX MEDLINE=85257431; PubMed=3848399;
 RA Evans B.A., Richards R.I.;
 RT "Genes for the alpha and gamma subunits of mouse nerve growth factor
 RT are contiguous";
 RL EMBO J. 4:133-138(1985).
 RN (3)
 RN SEQUENCE OF 25-261.
 RX MEDLINE=81264363; PubMed=7263706;
 RA Thomas K.A., Baglan N.C., Bradshaw R.A.;
 RT "The amino acid sequence of the gamma-subunit of mouse submaxillary
 RT gland 7 S nerve growth factor";
 RL J. Biol. Chem. 256:9156-9166(1981).
 RN (4)
 RN X-RAY CRYSTALLOGRAPHY (3.15 ANGSTROMS) OF 7S COMPLEX.
 RC STRAIN=swiss Webster; TISSUE=Submaxillary gland;
 RX MEDLINE=98035451; PubMed=9351801;
 RA Bax B., Blundell T.L., Murray-Rust J., McDonald N.O.;
 RT "Structure of mouse 7S NGF: a complex of nerve growth factor with
 RT four binding proteins";
 RL Structure 5:1275-1285(1997).
 CC -1- FUNCTION: 7S NGF ALPHA CHAIN STABILIZES THE 7S COMPLEX. THE BETA
 CC DIMER PROMOTES NERVE GROWTH. THE GAMMA CHAIN IS AN ARGININE-
 CC SPECIFIC PROTEASE; IT MAY ALSO HAVE PLASMINOGEN ACTIVATOR
 CC ACTIVITY, AS WELL AS MITOGENIC ACTIVITY FOR CHICK EMBRYO
 CC FIBROBLASTS.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage of Arg-1-Xaa bonds in
 CC small molecule substrates. Highly selective action to release
 CC kallidin (Lysyl-bradykinin) from kininogen involves hydrolysis of
 CC Met-1-Xaa or Leu-1-Xaa.
 CC -1- SUBUNIT: 7S NERVE GROWTH FACTOR IS COMPOSED OF TWO ALPHA CHAINS,
 CC A BETA DIMER COMPOSED OF IDENTICAL CHAINS, AND TWO GAMMA CHAINS.
 CC -1- MISCELLANEOUS: THIS PRECURSOR IS CLEAVED INTO SEGMENTS TO PRODUCE
 CC THE ACTIVE FORM OF THE GAMMA CHAIN, WHICH OCCURS NATURALLY AS

CC COMBINATIONS OF EITHER TWO OR THREE SEGMENTS HELD TOGETHER BY
 CC DISULFIDE BONDS: B1 + A OR B1 + C + B2.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, KALLIKREIN SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: X01389; CAA25645.1; -
 CC EMBL: X01798; CAA25928.1; -
 CC EMBL: X01799; CAA25930.1; -
 CC PIR: A00942; NMSG.
 CC PDB: 1SGF; 27-MAY-98.
 CC MEROPS: S01.170; -
 CC MGD: MGI:97322; Ngf.
 CC InterPro: IPR001314; Chymotrypsin.
 CC InterPro: IPR001254; Ser-protease_Try.
 CC Pfam: PF00089; trypsin; 1.
 CC PRINTS: PR00722; CHYMOTRYPSIN.
 CC SMART: SM00020; Tryp_Spc; 1.
 CC PROSITE: PS00240; TRYPsin_DOM; 1.
 CC PROSITE: PS00134; TRYPsin_HIS; 1.
 CC PROSITE: PS00135; TRYPsin_SER; 1.
 CC KMW: Hydrolyase; Serine protease; Glycoprotein; Multigene family; zymogen;
 CC signal; Growth factor; 3D-structure.
 CC FT SIGNAL 1 18 PROBABLE.
 CC FT PROPEP 19 24 ACTIVATION PEPTIDE.
 CC FT CHAIN 25 261 GLANDULAR KALLIKREIN K3.
 CC FT CHAIN 125 107 NERVE GROWTH FACTOR GAMMA CHAIN 1.
 CC FT CHAIN 112 261 NERVE GROWTH FACTOR GAMMA CHAIN 2.
 CC FT ACT_SITE 65 65 CHARGE RELAY SYSTEM.
 CC FT ACT_SITE 120 120 CHARGE RELAY SYSTEM.
 CC FT ACT_SITE 213 213 CHARGE RELAY SYSTEM.
 CC FT DISULFID 31 173
 CC FT DISULFID 50 66
 CC FT DISULFID 152 219
 CC FT DISULFID 184 198
 CC FT DISULFID 209 234
 CC FT CARBOHYD 102 102
 CC FT DOMAIN 25 107 SEGMENT B1.
 CC FT DOMAIN 112 261 SEGMENT A.
 CC FT DOMAIN 112 164 SEGMENT C.
 CC FT DOMAIN 165 261 SEGMENT B2.
 CC FT CONFLICT 108 111 MISSING (IN REF. 2).
 CC SQ SEQUENCE 261 AA; 28998 MW; 4870748E174AF7C8 CRC64;

Query Match 47.5%; Score 600.5; DB 1; Length 261;
 Best Local Similarity 44.7%; Pred. No. 4.7e-48;
 Matches 106; Conservative 46; Mismatches 78; Indels 7; Gaps 1;

QY 1 RIIFGECKRHSQPWQALFEKTRLLCGATLIPRLMLTAHCLKPRYIVHGOHNLQKE 60
 DB 24 RVIIGGCKEKNSSQPMWVAAYRYQYLOGVLLPNNVLAFAHCYDDMYKWLCKNNLFXD 83
 QY 61 ECGEOTRTATESRPHFGFNNSL-----DNKRDIMLVKASPVSTMAVRPLTLSS 113
 DB 84 EPPAQRHVLVSQSPHPYKPFELMKNRKRGDDSHNMLHLHSQPDITDGVKVDLPT 143
 QY 114 RCVYAGTSCILSGWGTSSPOLRLPRTLRCANTTIEHOKCEAAYPGNTDPMVCAVQOE 173
 DB 144 EEPKVGSTCLASGWSGSKPLPMEFPDLDQCVNHLHLSNECIRAYKREKVTIDMILCAGEMD 203
 QY 174 GGDSCQSGSGGGLPVNCSQIGIISWGDPCAITRRPGVTVKCYVYDIQETMKN 230
 DB 204 GGDTCYGDGSGGLDGVLOGITSWGHPGCAKTMNPAITKLTFTSWIKOTMARN 260

RESULT 14
 KLR3_PRANA

ID	KLKPR_PRANA	STANDARD:	PRT:	263 AA.
AC	P32824:			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	01-OCT-1993 (Rel. 27, Last sequence update)			
DT	01-NOV-1995 (Rel. 32, Last annotation update)			
DE	Glandular kallikrein, renal precursor (EC 3.4.21.35) (Tissue kallikrein).			
OS	Prorhynchus natalensis (African soft-furred rat) (Mastomys natalensis).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			
OC	Mastomys.			
OX	NCBI_TaxID=10112;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Salivary gland;			
RA	MEDLINE=944226702; PubMed=7909667;			
RT	Fahnestock M.;			
RT	"Characterization of kallikrein cDNAs from the African rodent Mastomys";			
RU	DNA Cell Biol. 13:293-300(1994).			
CC	- FUNCTION: GLANDULAR KALLIKREINS CLEAVE MET-LYS AND ARG-SER BONDS IN KININOGEN TO RELEASE LYS-BRADYKININ.			
CC	- CATALYTIC ACTIVITY: Preferential cleavage of Arg-I-xaa bonds in small molecule substrates. Highly selective action to release kallidin (lysyl-bradykinin) from kininogen involves hydrolysis of Met-I-xaa or Leu-I-xaa.			
CC	- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: X17352; CAA5232.1; -			
DR	PIR: S15686; S15686.			
DR	HSSP: P00757; 1SGF.			
DR	MEROPS: S01.160; -			
DR	InterPro: IPR001314; Chymotrypsin.			
DR	InterPro: IPR001254; Ser_protease_Try.			
DR	Pfam: PF00089; trypsin: 1			
DR	PRINTS: PR00722; CHYMOTRYPSIN.			
DR	SMART: SM00020; TRY-Spec: 1.			
DR	PROSITE: PS50240; TRYPsin_DOM: 1.			
DR	PROSITE: PS00134; TRYPsin_HIS: 1.			
DR	PROSITE: PS00135; TRYPsin_SER: 1.			
KW	Hydrolase; Serine protease; Glycoprotein; Multigene family; Zymogen; Signal.			
FT	SIGNAL	1	18	PROBABLE.
FT	PROPEP	19	24	ACTIVATION PEPTIDE (PROBABLE).
FT	CHAIN	25	263	GLANDULAR KALLIKREIN, RENAL.
FT	ACT_SITE	65	65	CHARGE RELAY SYSTEM.
FT	ACT_SITE	121	121	CHARGE RELAY SYSTEM.
FT	ACT_SITE	215	215	CHARGE RELAY SYSTEM.
FT	DISULFID	31	175	BY SIMILARITY.
FT	DISULFID	50	66	BY SIMILARITY.
FT	DISULFID	153	221	BY SIMILARITY.
FT	DISULFID	186	200	BY SIMILARITY.
FT	DISULFID	211	236	BY SIMILARITY.
FT	CARBOHYD	102	102	N-LINKED (GLCNAC...) (PROBABLE).
SO	SEQUENCE	263 AA;	29130 MW;	ABEB03B800337D5 CRC64;
QY	Query Match	47.2%;	Score 596.5;	DB 1; Length 263;
Db	Best Local Similarity	44.8%;	Pred. No. 1.1e-47;	
Matches	107;	Conservative	44;	Mismatches 79; Indels 9; Gaps 2;
QY	1	RIIKGECKPHSQPQAOALFEKTRILGCGATLAPLWMLTAAHCLPKRYIVHLGQNTLOKE	60	
Db	24	RIIGFNEKNKSNQPHAVAVYRFRARQGGGLVDAMWVLTAAHCHYDKXQVWLGKNNRED	83	
QY	61	EGCEGTRATSESPHPGNNSLPNKDH-----RNDIMLVKMAVPSITWAVRPLTL	112	

Db	84	ESBAGHQLSLKLRPHGFNMSLLKNDHTHPEDDYSNDMLVRLKRPALITDVKRIDL	143
Qy	113	SRGVYAGTISCLISGWSTIS-SPQLRLPHTLRKANITIIHOKECENAYPGNITDTMYCASV	171
Db	144	TEEPYVSRCLASGMOSTPTTEEFEXSHDLCQVYLELLSNVCARNAHTEKVDITMFCAGE	203
Qy	172	QEGKADSCGGDSGGPLVYCQSLQGITISWQDPCATIRKGVYTKYCKYDWMQETMKN	230
Db	204	MDGGKDTCCVDSGGPLICDGVLOGITISWGPPTCPALPNVGIYTKLEIVRSWKIDVANN	262
RESULT 15			
ID	TRY3_CHICK	STANDARD;	PRT; 248 AA.
AC	Q90629;		
DT	01-NOV-1997	(Rel. 35, Created)	
DT	01-NOV-1997	(Rel. 35, Last sequence update)	
DT	15-JUN-2002	(Rel. 41, Last annotation update)	
DE	Trypsin II-P29 precursor (EC 3.4.21.4).		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauirata; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
CC	Gallus.		
OX	NCBI_TaxID=9031;		
RP	[1]		
RN	SEQUENCE FROM N.A.		
RC	TISSUE=Pancreas;		
RX	MEDLINE=95251611; PubMed=7733885;		
RA	Wang K., Gan L., Lee I., Hood L.E.;		
RT	"Isolation and characterization of the chicken trypsinogen gene family.";		
RL	Blochem. J. 307:471-479(1995).		
CC	-I- CATALYTIC ACTIVITY: Preferential cleavage: Arg- -Xaa, Lys- -Xaa.		
CC	-I- SUBCELLULAR LOCATION: Extracellular.		
CC	-I- TISSUE SPECIFICITY: HIGH LEVELS ARE SEEN IN THE PANCREAS WHILE		
CC	LOWER LEVELS ARE FOUND IN THE LIVER, SPLEEN AND THYMUS.		
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/isb-sib.ch).		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
CC	EMBL: U15157; AAA79914.1; -.		
DR	HSSP: P00763; IDPO.		
DR	MEROPS: S01.151; -.		
DR	InterPro: IPR001314; Chymotrypsin.		
DR	InterPro: IPR001254; Ser_protease_Try.		
DR	Pfam: PF00089; trypsin; 1.		
DR	PRINTS: PR00722; CHYMOTRYPSIN.		
DR	SMART: SM00020; Tryp.Spec. 1.		
DR	PROSITE: PS50240; TRYPSIN_DOM; 1.		
DR	PROSITE: PS00134; TRYPSIN_HIS; 1.		
DR	PROSITE: PS00135; TRYPSIN_SER; 1.		
KW	Hydrolase; Serine protease; Digestion; Pancreas; zymogen; Signal;		
	Multigene family.		
FT	SIGNAL	1	16
FT	PROPEP	17	25
FT	CHAIN	26	248
FT	ACT_SITE	65	65
FT	ACT_SITE	109	109
FT	ACT_SITE	202	202
FT	DISULFID	32	162
FT	DISULFID	50	66
FT	DISULFID	134	235
FT	DISULFID	141	208
FT	DISULFID	173	187
FT	DISULFID	198	222
FT	STATE	196	196
			REQUIRED FOR SPECIFICITY (BY SIMILARITY).

```

SQ SEQUENCE 248 AA: 26622 MW: E5F16B07622B588E CRC64:
Query Match 47.2%; Score 596; DB 1; Length 248;
Best Local Similarity 49.3%; Pred. No. 1.2e-47;
Matches 112; Conservative 36; Mismatches 73; Indels 6; Gaps 3;

QY 1 RIRKEPCKPHSQPMQALFEKTRLLCGATLAPRWLTAAACLPRIYVHGHONLKE 60
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 25 KIVGGYTCEHSPYQVSL-NSGYHFCGSLINSQVLSAAHCYSRIQVRLGEYNIDVQ 83
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 61 EGCQETRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPVSTWMAVRPLTSSRCVTAGT 120
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 84 EDSVEYVRSSSVIIRHPKXSSITLN-----NDIMLIKLASVRESADIQPIALSSCAKAGT 139
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 121 SCLISGWSSTSPQRLPHTLRKANITITIEHOKCENAVPGNTITWVCASVQEGKDSQ 180
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 140 ECLISGWSGNTLSNGYNPELLQCLNAPILSDQCEQAYPGDITSNMIGVFLGEGKDSQ 199
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 181 GDSGSPVNCOSLOGIISMGDPACATITRPGYTVCKYVDMIOETM 227
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 200 GDSGSPVNCNGLQIVSMWIG-CALKGYPGYTVCKYVDMIOETI 245
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 16
TRY2_RAT STANDARD: PRT: 246 AA.
ID TRY2_RAT
AC P00763;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin II, anionic precursor (EC 3.4.21.4) (Pretrypsinogen II).
GN TRIZ.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85054880; PubMed=6094547;
RA Craik C.S., Choo Q.L., Swift G.H., Quinto C., McDonald R.J.,
RA Rutter W.J.;
RT "Structure of two related rat pancreatic trypsin genes.";
RL J. Biol. Chem. 259:14255-14264(1984).
RN [2]
RP SEQUENCE OF 9-246 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Pancreas;
RX MEDLINE=82265624; PubMed=6896710;
RA McDonald R.J., Stary S.J., Swift G.H.;
RT "Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide
RT sequences of the cloned cDNAs.";
RL J. Biol. Chem. 257:9724-9732(1982).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.59 ANGSTROMS).
RX MEDLINE=91351998; PubMed=1881877;
RA Earnest T., Fauman E., Craik C.S., Stroud R.;
RT "1.59-A structure of trypsin at 120 K: comparison of low temperature
RT and room temperature structures.";
RL Proteins 10:171-187(1991).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=96214506; PubMed=8634241;
RA Brien L.S., Willett W.S., Craik C.S., Fletterick R.J.;
RT "X-ray structures of a designed binding site in trypsin show metal-
RT dependent geometry.";
RL Biochemistry 35:5999-6009(1996).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V01274; CAA24581.1; -.
DR EMBL: L00131; AAA98517.1; -.
DR EMBL: L00130; AAA98517.1; JOINED.
DR PIR: A22657; TRIZ2.
DR PDB: 1ANB; 01-APR-97.
DR PDB: 1ANB; 01-APR-97.
DR PDB: 1AND; 01-APR-97.
DR PDB: 1ANE; 01-APR-97.
DR PDB: 1AMH; 24-DEC-97.
DR PDB: 1SLU; 11-JUL-96.
DR PDB: 1SLV; 11-JUL-96.
DR PDB: 1SLW; 11-JUL-96.
DR PDB: 1SLX; 11-JUL-96.
DR PDB: 1DPO; 07-JUL-97.
DR PDB: 3TGT; 23-DEC-98.
DR PDB: 3TGT; 23-DEC-98.
DR MEROPS: S01.258; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser.protease_Try.
DR Pfam: PF00089; trypsin_1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRYP-Spc. 1.
DR PROSITE: PS50240; TRYP-SIN_DOM. 1.
DR PROSITE: PS00134; TRYP-SIN_HIS. 1.
DR PROSITE: PS00135; TRYP-SIN_SER. 1.
KW Hydrolase, Serine protease, Digestion, Pancreas, Zymogen, Signal;
KW Multigene family, 3d-structure.
FT SIGNAL 1 15
FT PROPEP 16 23
FT CHAIN 24 246
FT ACT_SITE 63 63
FT ACT_SITE 107 107
FT ACT_SITE 200 200
FT DISULFID 30 160
FT DISULFID 48 64
FT DISULFID 132 233
FT DISULFID 139 206
FT DISULFID 171 185
FT DISULFID 196 220
FT SITE 194 194
FT SITE 194 194
FT CONFLICT 84 84
FT CONFLICT 88 88
SQ SEQUENCE 246 AA: 26228 MW: ABD3630809AEE606 CRC64;

Query Match 47.0%; Score 593; DB 1; Length 246;
Best Local Similarity 49.1%; Pred. No. 2.2e-47;
Matches 113; Conservative 33; Mismatches 78; Indels 6; Gaps 3;

QY 1 RIRKEPCKPHSQPMQALFEKTRLLCGATLAPRWLTAAACLPRIYVHGHONLKE 60
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 23 KIVGGYTCEHSPYQVSL-NSGYHFCGSLINDQVWVSAACHYSRIQVRLGEYHINVL 81
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 61 EGCQETRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPVSTWMAVRPLTSSRCVTAGT 120
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 82 EGNQGFVNAKTIKHPNPKRLTN-----NDIMLIKLSVSKLNARVATVALPSSCAKAGT 137
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 121 SCLISGWSSTSPQRLPHTLRKANITITIEHOKCENAVPGNTITWVCASVQEGKDSQ 180
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 138 OCLISGWSGNTLSNGYNPELLQCLNAPILSDQCEQAYPGDITSNMIGVFLGEGKDSQ 197
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 181 GDSGSPVNCOSLOGIISMGDPACATITRPGYTVCKYVDMIOETMKN 230
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 198 GDSGSPVNCNGLQIVSMWIG-CALPDNPGYTVCKYVDMIOETIAN 246
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

Search completed: December 30, 2002, 14:50:22
Job time : 13 secs

> 0 <
0110 Intelligenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-856-320a-2.res made by bobdren on Mon 30 Dec 102 15:06:50-PST.

Query sequence being compared: US-09-856-320A-2 (54-282)
Number of sequences searched: 1
Number of scores above cutoff: 1

Results of the initial comparison of US-09-856-320A-2 (54-282) with:
File : baal1306.pep

```

100-
N -
U - 50-
M -
B -
E -
R -
O -
F - 10-
S -
E - 5-
Q -
U -
N -
C -
E -
S - 0-
SCORE 0 14 27 41 54 68 81 95 108 122
STDDEV

```

PARAMETERS

Similarity matrix PAM-150 K-tuple 1
Threshold level of sim. 16%
Mismatch penalty 1 Joining penalty 20
Gap penalty 5.00 Window size 229
Gap size penalty 0.05
Cutoff score 1
Randomization group 0

SEARCH STATISTICS

Scores: Mean 122 Median 0 Standard Deviation 0.00
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00
Number of residues: 244
Number of sequences searched: 1
Number of scores above cutoff: 1

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Opt. Frame
1. baal1306	TOIG of: baal1306	check: 178	244	122 162 0.00 0

1. US-09-856-320A-2 (54-282)
baal1306 TOIG of: baal1306 check: 1789 from: 1 to: 244

TOIG of: baal1306 check: 1789 from: 1 to: 244

BAAL1306 GB:D78203 neurosin [Homo sapiens] (ver 1)

BAAL1306 Length: 244 December 30, 2002 14:51 Type: P Check: 1789 ..

Initial Score = 122 Optimized Score = 162 Significance = 0.00
Residue Identity = 44% Matches = 103 Mismatches = 106
Gaps = 6 Conservative Substitutions = 14

baal1306 KKIMVLSLIAANAEEGNKLVHGSPCDKTSHPYQALVTSGLHCGVLIHPLWVLTAAHCKRNQVFLS
10 20 X 30 40 50
10 20 X 30 40 50 60 70

OHNLQKEEGCEQFTRATSPFPHPGPNNSLPNKHDRHNDIMLVKMASPVITWAVRPLTSSRCVAGTSCLS
111
KHNLQRESSQEQSSVAVYIHPDIYDAA-----SHDQIMLRAPARLSEIOLPLEROCSSANTTSCHL
80 90 100 110 120 130 140

GMGSTRSPQLRLPHTLRCAANTITIEHQKCEANAYPGNITDTWYCASVQEGKDCGDSGGLVYCNQSLQGIT
111
GMGKTADGD--FPDTIQCAIYHILVSRECEHAYRQITQNNLCMEDEYRGKDCGDSGGLVYCNQSLQGIT
150 160 170 180 190 200 210

200 210 220 230 240 X
SMGDDPCAIRRRPGVYTKVCKYVDIMQETMKN
111
SMGNIPCGSKERPGVYTNVCRTNMIOKTIQAK
220 230 240 X

THIS PAGE BLANK (USPTO)